

Special States Demand a Force for *the* Observer

L. S. Schulman^{1,2}

Received: 25 February 2016 / Accepted: 9 June 2016 / Published online: 28 July 2016
© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract The “special state” understanding of the measurement process is presented, namely there is no “measurement process,” only unitary time evolution. However, in contrast to the many worlds interpretation, there is only one world. How this can be accomplished and how statistical mechanics is changed as a result are also discussed. The focus though is on experimental tests of this theory and the in-principle realization that this can give rise to feasible experimental tests. Those tests rely on the particular feature of having only one world, so that any change in the wave function must have a proximate cause, and it is the detection of that cause that constitutes the test. In a companion article there is further exploration concerning the details of the test. In addition, in the present article, the special state theory is extended theoretically through evidence of the uniqueness of the Cauchy distribution as well as explicit recognition of the role of entanglement.

Keywords Special states · Arrow of time · Foundational experiments

1 Introduction

Many interpretations of quantum mechanics exist, all of which agree on experimental results. This is why they are called *interpretations*. Other understandings of the measurement process may differ and sometimes offer the possibility of an experimental check. In this article I describe a theory that differs in its experimental predictions from

✉ L. S. Schulman
schulman@clarkson.edu

¹ Max Planck Institute for the Physics of Complex Systems, Nothnitzer Str 38, 10187 Dresden, Germany

² Physics Department, Clarkson University, 5820, Potsdam, NY 13699, USA

all others that I know of. Like the Many Worlds Interpretation it allows only unitary time evolution, but unlike that theory, has only one world. This suffices to create an experimental test.

The experimental prediction to which I refer is the absence of source free changes in angular momentum. Thus in either the many worlds interpretation (MWI) or the Copenhagen interpretation (in which I include the variety of ensemble theories) it is permitted for a spin prepared (say) in an eigenstate of J_x to be measured in an eigenstate of J_z . Your idea of which way the angular momentum vector is pointing has changed, but there is no need for any force to have been applied (I go into this in detail below). For the special state theory, there *must* be a force applied to the spin to reorient it. Observing or not observing that force is the test.

In Sect. 2 I discuss quantum aspects, how it is possible to have *only* pure unitary (time) evolution and nevertheless have but a single “world.” Following that, Sect. 3 deals with related issues of statistical mechanics. This is relevant, since it is clear that to have both pure unitary evolution and only one world *something* wildly un-intuitive must be happening. This strangeness will be seen to be in the realm of statistical mechanics. Some of these topics are taken up at greater length in [1].

Finally in Sect. 5 I show that for the special state theory the changes in direction of one’s perception of angular momentum require forces, while for the MWI (etc.), they do not. Sect. 6 is a discussion.

The appendices are mainly concerned with technical matters. Of note is Appendix 2, where I present strong evidence for the appearance of the Cauchy distribution in the theory. (See Sect. 4 for the significance of this distribution.) The material in Appendix 2 represents an important step beyond what appears in [1]. In that reference I found that the Cauchy distribution provided a solution, but there was no proof that the second moment needed to be infinite. In the new work the drop-off of the Cauchy distribution stands out as superior to other options including other long-tailed distributions that lack a second moment.

2 Quantum Aspects

According to the special state theory, the only thing that ever happens is pure unitary quantum time evolution. If one has a wave function ψ at time-0, and if H is its Hamiltonian, then at time- t the wave function is $\exp(-iHt/\hbar)\psi$. “Measurement” involves no other dynamics. So far this sounds like the many-worlds interpretation. But I inject another feature: there is only one world. How can this be? Let us look at a very small world. Suppose there is a single 2-state system in contact with a heat bath of bosons [2]. Initially the system is in its excited state and the bosons serve two purposes: their coupling induces the decay and they constitute the measurement device that notes the decay. This model of decay and measurement misses much of the real world, like “registering” the measurement, i.e., making sure the system does not return to its excited state (irreversibility). These features are assumed to result from parts of the system

that I do not model. I use the spin boson model with a single boson, for which I can take the Hamiltonian¹ to be

$$H = \frac{\varepsilon}{2} (1 + \sigma_z) + \omega a^\dagger a + \beta \sigma_x (a^\dagger + a). \quad (1)$$

The Pauli spin matrices are the operators for the 2-state (spin) system, a and a^\dagger are the boson operators and ε , β and ω are parameters.

Suppose this system is started with spin up, $\psi_{\text{spin}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, with the oscillator state unspecified. For the parameters given in footnote 2, at time-0.15 there is about a 50 % probability of decay, namely if one would trace out over the oscillator states the density matrix for the spin would be half-half. However, there are two states that I wish to focus on. These are particular initial states of the oscillator, and their probability distributions are given in Fig. 1. The state whose probability distribution is shown in Fig. 1a (and whose phases are also fixed, but not shown) has the property that at time-0.15 it has not decayed at all, despite the fact that a random or average state would, at that time, be in a superposition of up and down spin states, with about half its probability in each state. Similarly, the state whose probabilities are shown in Fig. 1b has decayed (almost)³ completely at this time, again despite the behavior of a typical or average state.

This does not mean that the non-decay state must never change its spin-occupation values and certainly the decay state had to change, since at time-0 it had no amplitude in the decayed state. For reference, the probability that each of these states is in the “up” state is shown as a function of time in Fig. 2. As is clear, the decay or non-decay criterion is *only* applied at time-0.15. At other times the states can be anything.

What good are these states and how did I find them? The second question is addressed in Appendix 1 and is technical. The answer to the first question is the main idea for avoiding *many* worlds while holding on to unitary time evolution.

Suppose our Schrödinger cat is placed in a chamber with the usual vial of poison whose dispersal is governed by the spin state of a spin-boson system of the sort just discussed. The latter system is also in the chamber and the entire setup isolated. It is opened—its isolation ceases and the irreversible “registration” is due to the observer—at time-0.15 and one looks to see if the cat is dead or alive. The usual problem is that there is positive probability for both options. However, there is *no* problem if the initial state of the oscillator is one of the “special” ones I have been describing. For the non-decay state there is a living cat, for the decay state it is dead. This is accomplished

¹ The full spin boson model is generally taken to have Hamiltonian $H = (\varepsilon/2)(1 + \sigma_z) + \Delta \sigma_x + \sum_k \omega_k a_k^\dagger a_k + \sigma_z \sum_k \beta_k (a_k^\dagger + a_k) + \sum_k \beta_k^2 / \omega_k$, but we will take a much simpler version for our example, as given in Eq. (1).

² The parameters for the spin boson model of Eq. (1) are $\varepsilon = 0.5$, $\omega = 0.1$, and $\beta = 0.6$. The oscillator was cut off and altogether 250 states were considered. This led to an error in the commutator of a and a^\dagger in the 250th diagonal term, but not elsewhere. To reduce the cut off effect, only states with relatively small probability in the highest levels were considered.

³ Neither the non-decay nor the decay is perfect. For the case shown, the probability of decay for the “non-decay” state is about 6.1×10^{-4} , while the probability of non-decay of the “decay” state is about 8.1×10^{-5} . I comment on this near the end of Sect. 3.

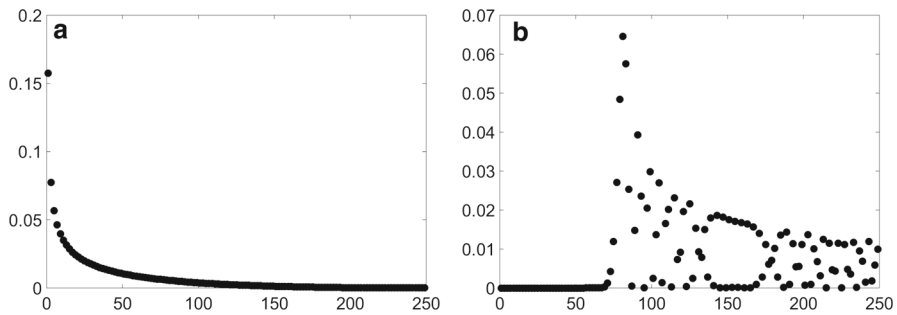


Fig. 1 **a** shows the probability of excitation of various oscillator states that contribute to the non-decay state. Only shown are even oscillator states, as the total amplitude in this case for the odd states is computed to be about 10^{-27} and is due to numerical error.⁴ Phases of the states are not shown, but are also fixed by the non-decay condition. In image **(b)** are shown the probabilities for the state that decays; in this case only even *oscillator states* are shown (again the wrong parity states are due to numerical error). As in image **(a)**, the phases, though not shown are crucial to the “special” nature of the state. *Both images* represent the probabilities *at time-0*. Note that although in this case both special states only involve even states, in general special states can be of either parity. [On the other hand, for other Hamiltonians or projections, the special states need not be eigenstates of H (see footnote 4)].

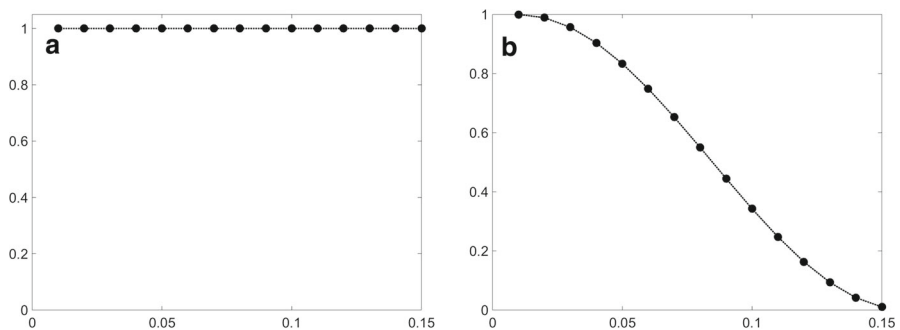


Fig. 2 Survival history of the non-decay and decay states, where “survival” is the probability that the system is in the “up” state at the indicated time (*horizontal axis*). Image **(a)** shows this probability for the non-decay state, which for this example was a “non-decay” state at all intermediate times, but in general need not be. Similarly, image **(b)** shows the corresponding probability for the state found to be (essentially) fully decayed at time-0.15. At both earlier and later times it need not be fully decayed.

with no black magic; it is the result of unitary time evolution from the specified initial conditions.

That is the main idea of the special state theory: no macroscopic superpositions because the initial conditions are special—always.⁵ I mention also that there is no

⁴ For the Hamiltonian of Eq. (1) there is a constant of the motion, conventionally called “parity.” For the single oscillator this is $\text{parity} \equiv \Pi = (-1)^{a^\dagger a} \sigma_z$, and $[H, \Pi] = 0$. (See Appendix 1 for notation in the following discussion.) Our “ B ” also commutes with Π because the projections involved (P and Q) themselves commute with Π , since they are functions of the operator σ_z . Hence the eigenstates of B also can be sorted by their parity. In general, if the projections do not commute with particular symmetries of the Hamiltonian there would be no need for B and H to have common spectrum.

⁵ Note also that this enforces a degree of determinism that may elicit extreme discomfort. For example, my special state for the spin-boson system put into the cat’s chamber is special for time 0.15, not for

entanglement. At time-0.15 the spin state is wholly in one state or the other and a trace over its coordinates would leave the oscillator state unchanged; and vice-versa.

The next—obvious—question is, why should Nature arrange to have a “special” state as the initial conditions for every situation where a potential split into many worlds occurs? For this I do not have an answer, except to say that it is the conclusion I am driven to by insisting that no magic dynamics occurs in the measurement process and there is only one world. What I can offer though is perspective. How strange is it for there to be particular, non-random, initial conditions? For this I turn to the next Section.

3 Boundary Conditions and the Arrow of Time

When making predictions I assume the initial conditions are random. As discussed in [1] this is equivalent to the usual arrow of time. But this assumption has never been verified experimentally; its main virtue is that answers computed with this assumption agree with experiment—no mean feat, but not a definitive proof, as I will shortly demonstrate.

This assumption is closely related to another that occurs in statistical mechanics, the ergodic hypothesis. As discussed in [3] this lies at the foundations of statistical mechanics but has never been established⁶ and is very likely false. Textbook authors (as described in [3]) have struggled with this problem in arguments leading to the justification, for example, of adopting the thermal state ($\rho \sim \exp(-\beta H)$). But the hypothesis is generally accepted for the reason mentioned earlier: agreement with experience—so far.

To show that the assumption of random initial conditions is unnecessary, I use the cat map [4], which is known to be mixing (and thus ergodic). This is a map of the unit square (with coordinates x, y) into itself

$$\left. \begin{array}{l} x' \equiv x + y \\ y' \equiv x + 2y \end{array} \right\} \bmod 1 \quad \text{or} \quad \begin{pmatrix} x' \\ y' \end{pmatrix} = M \begin{pmatrix} x \\ y \end{pmatrix} \bmod 1 \quad \text{with} \quad M \equiv \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}. \quad (2)$$

Our system is an ideal gas consisting of N points in the unit square, each moving in discrete time under the cat map. For example, in Fig. 3 I show what happens with a collection of $N = 500$ points initially satisfying $0.5 \leq x \leq 0.6$ and $0.5 \leq y \leq 0.6$. Clearly this mechanical system is headed for chaos. To get a quantitative idea of “chaos” I define an information entropy using a coarse graining. As grains I take the

Footnote 5 continued

another time. It is coordinated with the *fact* that the observer actually opens the chamber at that time. This observer may think the opening time is arbitrary, but it is already built into the state of the universe that the chamber will be opened at that time. For some interpretations of the concept of free will this would deny that possibility (but there are *many* interpretations). This level of predictability and determinism is natural with the two-time boundary conditions that are introduced in the next section, but for some, the conclusions may be shocking.

⁶ Proofs of ergodicity in the realm of mathematics of course do exist, but they are of little relevance for the uses to which physicists put this hypothesis. The dynamics used in those demonstrations is artificial and even more important, the time scales for true multiparticle systems are enormous, well beyond the lifetime of the universe.

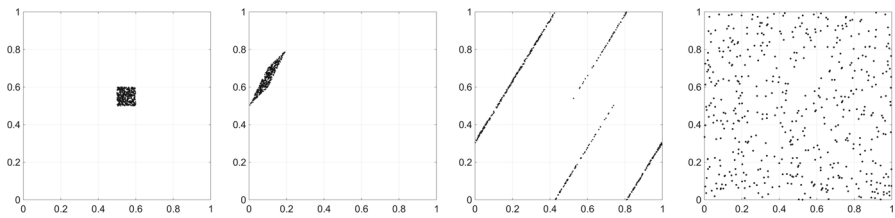
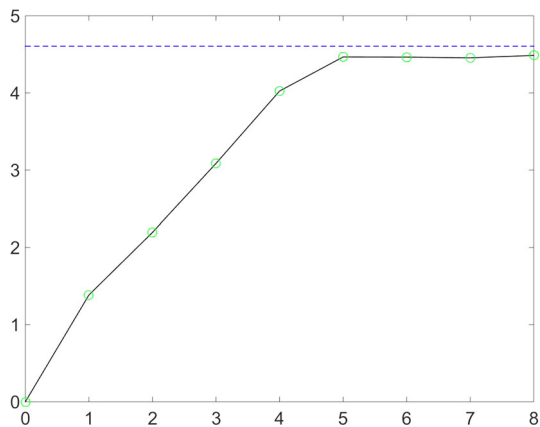


Fig. 3 500 points are started in the square $0.5 \leq x \leq 0.6$ and $0.5 \leq y \leq 0.6$, but are otherwise randomly selected. After one time step under the cat map, Eq. (2), they have become the parallelogram in the next figure. They have stretched in the direction of the eigenvector with larger eigenvalue of M (Eq. (2)) and correspondingly shrunk along the other eigenvector. The product of the eigenvalues is 1 because $\det M = 1$. The dynamics thus satisfies Liouville's theorem, in which area is preserved in phase space. By time 3 (the next figure) the mod 1 action coupled with the stretching has begun to pull the points apart and by time 7 (the last figure) nothing is recognizable.

Fig. 4 Entropy, as defined in Eq. (3), as a function of time for the simulation of Fig. 3. Note that equilibration sets in at about time-5. The dashed figure is the maximum entropy for this coarse graining, namely $\log 100$. It is not attained because of fluctuations due to the finite number of points.



100 $1/10$ by $1/10$ squares contained in the unit square, and only count the number of points in each grain. In effect I assume the observer can only determine which square a point is in, not the point's exact coordinates. With this coarse graining the entropy is

$$S = - \sum p_k \log p_k, \quad p_k \equiv n_k/N, \quad (3)$$

where k runs over the coarse grains and n_k is the number of points in grain- k . The behavior of the entropy for the points in Fig. 3 is shown in Fig. 4

This is what one might call normal behavior: the initial points were selected randomly within the given grain and the entropy increases monotonically until close to equilibrium, after which it fluctuates in a predictable way.⁷ But let me display another simulation. As I will explain in a moment, the initial points were *not* selected randomly, but for the first time steps it certainly will look that way. The simulation runs

⁷ Finding the average deviation from maximum entropy is straightforward for given parameters (number of grains, number of points), although the calculation involves non-uniform asymptotics. See [1], p. 40 and Exercise 2.2.2 (but beware of typos).

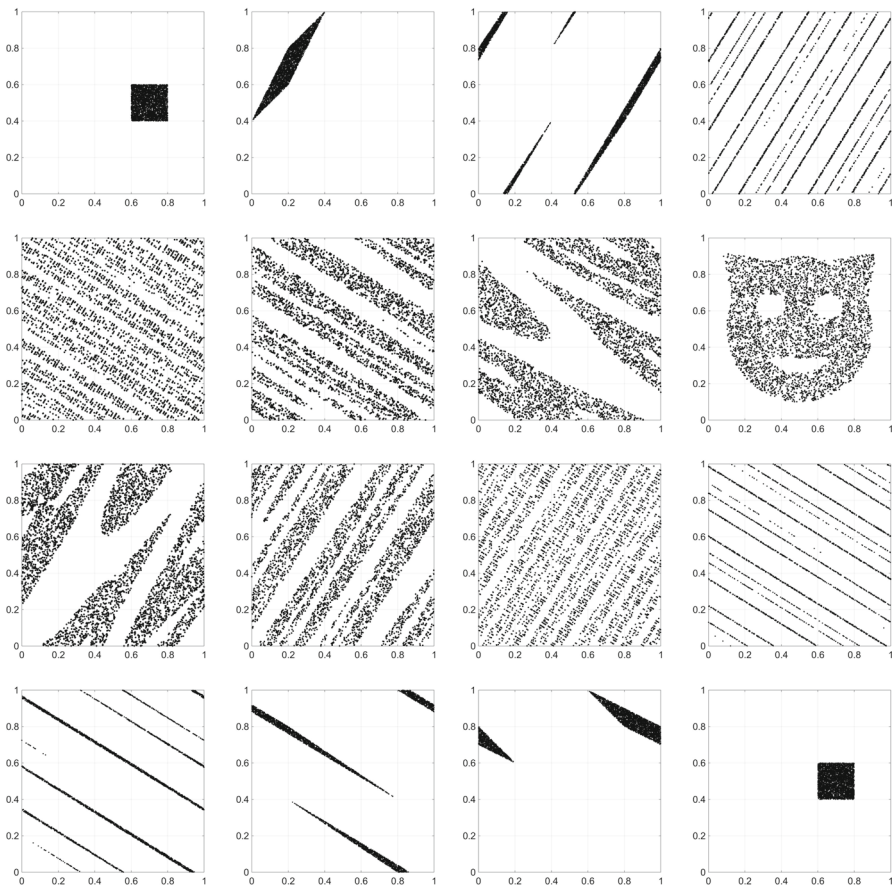
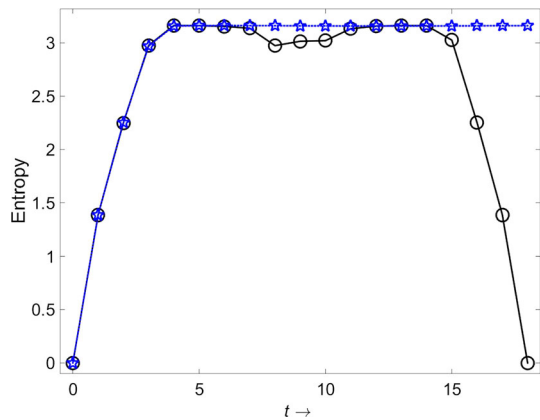


Fig. 5 The times are as follows: row-1: [0, 1, 2, 4]; row-2: [6, 7, 8, 9]; row-3: [10, 11, 12, 14]; row-4: [15, 16, 17, 18]. These points all evolve under the cat map and satisfy boundary conditions at times 0, 9 and 18. Size of the point position marker varies with the image, for better visibility.

for 18 times steps away from the initial square and is shown in Fig. 5. The sequence of images should be read left-to-right and row-by-row. There are 4000 points and most time steps are illustrated. I stress, *every single point in this simulation evolved by pure cat map dynamics*. So how did I get it to give me these strange images? Actually it was easy. I randomly occupied the first little square ($0.6 \leq x \leq 0.8$ and $0.4 \leq y \leq 0.6$) at time-0 with about 30×4000 points. Then I imposed two conditions on the points. First, at time-18 they needed to occupy a different little square (but of the same size as the original). This cut the number of acceptable points by a factor 25. Then I also required that at time-9 the points arranged themselves in the figure of a cat. This cut things down a bit more, so I was left with 4000 points that satisfied all three conditions. This solved a 3-time boundary value problem. It was trivial in this case since the points were non-interacting; simply removing a point did not affect the others. It should be pointed out that solving a multiple time boundary value problem for *interacting* particles, even two of them, can be quite challenging. But for the point I wish to make, an ideal gas does the job.

Fig. 6 The circles and solid line represent the entropy as a function of time for the simulation shown in Fig. 5. Note that it drops a bit at time-9 and then at time-18 plummets back to 0, since all points are again in a single grain. The other curve, marked by stars, is the entropy as a function of time for 4000 points having the same initial condition, but with no other constraints.



To get a message from this demonstration I turn to entropy. The coarse graining is a bit more coarse than before: grain sizes are $1/5$ by $1/5$. For each of the configurations shown the entropy was computed and is graphed as the circles in Fig. 6. A second curve, with star markers, is also shown in that figure. It is the entropy for 4000 points initially starting in the same coarse grain, but having no additional requirements. (Like Fig. 4, it corresponds to “normal” initial conditions.) Compare these two curves. The point I wish to make is that for times prior to about 7 you **cannot** tell the difference.

The entropy graph makes the point quantitatively, but qualitatively, if you compare Figs. 3 and 5, it is clear that aside from slightly different initial conditions, there is a great similarity in the initial relaxation. One can be quantitative in other ways also, discussing relaxation times and such, but the issue to be stressed is that with fewer than 4 % of available points the relaxation, almost to the time of the constraint, looks entirely normal.

My message is simple: I may have future constraints, but I would not know about it. The arrow of time does not (necessarily) point as fixedly as one might have supposed.

Another way to say this is to observe that the points in the simulation that yielded Fig. 5 were not random but had a *cryptic* constraint, that is a constraint that was difficult (or for the macroscopic world, impossible) to discern, but which nevertheless plays an important role as the dynamics unfolds.

My objective in this presentation is to pave the way for the idea that there could be other cryptic constraints in the world. In particular, not every imaginable state occurs in Nature; only those which, in my terminology, are special. This is surely a severe restriction, but I have made the point that the restriction may be invisible. The kind of restriction that I find most palatable is a two-time boundary condition. This surely contradicts the usual arrow of time, but as demonstrated, its effect may not be noticed except close to the boundaries, which may be well-separated.⁸

⁸ Thomas Gold introduced the idea that the thermodynamic arrow of time is a consequence of the expanding universe [8]. Some reviled this idea claiming the absurdity of an opposite arrow if our universe has a big crunch in its future. Even some respected scientists failed to appreciate how a two-time boundary value problem deals with this issue. See also [12].

What kind of two-time boundary condition could select special states? First consider initial conditions. As Wald [5] has pointed out, in the early universe the entropy was low, for unknown reasons. I will go a step further: perhaps the von Neumann entropy was also low, in other words there was little or no entanglement. In this speculative mode I will further imagine that our entire cosmology is roughly time symmetric. This idea is not popular today due to the discovery of accelerating expansion. But the phenomenon is poorly understood and there have been suggestions of a periodic cosmology despite the acceleration (a far from comprehensive sample is [6,7]), so I will take liberties in my speculation. One additional component enters this line of thought, namely the connection first suggested by Gold [8], relating the arrow of time to an expanding universe. One then expects that under contraction the arrow will be reversed. Recalling the connection between boundary conditions and the arrow of time, one can now enunciate a possible boundary condition that would demand special states: no entanglement at the beginning, no entanglement at the end.⁹

Consider then a Schrödinger cat. At the end of the experiment—and again demanding only unitary time evolution—in the MWI there is a portion of the wave function with a dead cat and portion of the wave function having support on a macroscopic state recognized as a living cat. The dead one is buried, the living one perhaps becomes involved in another experiment, sending cats to Mars. But how can these portions of wave function be recombined coherently, as they would need to be if there is a no-entanglement demand in our future. It would take tremendous coordination to accomplish this coherently. Having a special state is *also* an unlikely way to avoid entanglement, but it is much *less* unlikely than recombining after a superposition of macroscopically distinct states has formed.

There are many caveats in the above argument. The first is that maybe the need for special states has nothing to do with cosmology. It may be that special states are indeed the way to reconcile quantum and statistical mechanics, but the argument just given is wrong or irrelevant. Next, it is possible that future boundary conditions might obtain even in an ever-expanding universe. Then there are technical matters. What is the role of identical particles? Perhaps you do not need to recombine the cat, since its electrons and other constituents are all identical, you may be able to recombine local portions of the wave function with particles nearby. For these and other questions, the short answer is, I don't know.

But with the two-time boundary rationale one can approach another issue, namely what about the small amount of left over wave function. As I state in footnote 3, the special state for decay has a small but non-zero probability of non-decay (in the example given it is 8.1×10^{-5}), similarly for the non-decay special state. In the context of a boundary value problem one does not need perfection. The measure of possible error in “specializing” is given by the tolerance of that boundary value problem.

There are several comments to be made that reflect on the plausibility of these ideas. First there are the numbers. In my computer modeling I found errors of order 8.1×10^{-5} using a Hilbert space of dimension 500. The actual dimension of physical

⁹ Of course there is entanglement every time a bound state forms. The assumption is that this is also accomplished by means of special states. The electron and proton (say) definitely come together or definitely do not. As usual this can only be accomplished with the aid of other degrees of freedom.

spaces boggles the imagination. Using known formulas for entropy, one mole of neon in 1 cubic meter at room temperature and pressure possesses [3] on the order of $\mathcal{M} = \exp(S/k_B) \simeq 10^{1.3 \times 10^{25}}$ states, i.e., dimensions in Hilbert space. So for any macroscopic apparatus one can expect far smaller errors than I obtained in my example. In a similar vein, reflecting on the total number of states in the universe, it becomes less implausible that only a tiny subset (the special states) can nevertheless contain many states.¹⁰ Another point, which may be surprising, is that for the right kind of wave functions, there is a cessation of entanglement in their collisions [9]. If you scatter particles of unequal mass having Gaussian wave packets off one another they rapidly adjust the packet width so that there is no further entanglement.

Finally there is the issue of recovering the Born probabilities. If every experiment involves interaction with apparatus and special states, why is it that probabilities can be calculated using only the wave function of the system being studied? This is our next subject.

4 Recovering Standard Probabilities

A fair coin has a 50 % chance of landing heads, 50 % tails. In effect, the phase space of your body and the motions needed to flip the coin are divided into two sets, one of which gives heads, one tails. These should be closely interwoven so that you have difficulty controlling the outcome. So is this a property of the coin or of you?

Something along these lines is what I claim occurs during a quantum measurement. The Born rule says, look only at the wave function. But I am saying that the space of special states of the apparatus breaks up in the way the phase space of the coin flipper does. The special states are a small subset (subspace, actually) of the entire Hilbert space, but their relative size (dimension) is proportional to the probabilities of the various outcomes. In other words, quantum probability is like its classical counterpart: in each instance the result is determined, but one would need microscopic precision to know the outcome. So one uses probability. The claim therefore is that the dimension of the space of special states for each outcome is proportional to the square of the wave function amplitude for that outcome.

This is an assertion that I do not know how to check. The few cases where I have an analytic handle on special states (see [1]) I do not consider typical, and where I have

¹⁰ The ideas of determinism (see footnote 5) and two-time boundary conditions can have implications outside of the quantum issues we have devoted this article to. As is well-known, the evidence for the existence of black holes, while convincing (e.g., [13]), does not exclude the possibility that there are compact objects, highly dense, but that they simply have not passed to the black hole stage. Why not? No reason has been given and it would seem to be highly contrived to assume that peculiar dynamical features have just managed to prevent black hole formation under accepted notions of general relativity. However, once you have black holes in the universe other problems arise, information issues and firewalls [14]. Given the speculative nature of the boundary conditions I have proposed (“no entanglement at the beginning, no entanglement at the end”) one can continue the speculation and demand that the boundary conditions also require that no black holes form, despite their dynamical possibility (just as I am claiming that macroscopic superpositions—Schrödinger cats—do not form, despite their dynamical possibility). This may also have bearing on the question that Wald [5] poses: why did the universe not begin in a maximum entropy state, i.e., a black hole? On the other hand, my proposal can be viewed simply as a rephrasing of his question: why have these boundary conditions? It is questions like these that suggest humility for *homo sapiens*.

numerical results I believe the spaces are much too small. As a result, I have taken a defensive position on this point, to see if the assertion can be *disproved*.

I focus on the simplest quantum measurement problem, determining the state of a two-level system. I consider a Stern-Gerlach apparatus for which the initial spin state is

$$u_\theta = e^{i\theta\sigma_x/2} \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \quad (4)$$

There are of course many other degrees of freedom in this problem. Foremost is the position of the spin (riding perhaps on a K atom); then there is the macroscopic magnet, macroscopic screen, and much, much more. Consistent with the desire to show impossibility I grant that there are special states to do the job, different ones for different angles, θ , of the spin orientation. I can make a guess at the nature of the special state—if it exists—by looking for the least unlikely way to get the usual Stern-Gerlach result. The special state might unite the spatial wave function of the K with diverging position coordinates after it has emerged from the inhomogeneous magnetic field. This seems to me less likely—involving more degrees of freedom doing atypical things—than to have θ turned to 0 or π *before* it enters the field.¹¹ So we'll assume that's what happens; perhaps what one takes to be a stray magnetic field provides just the right force to bring the spin to 0 or π prior to its entry into the inhomogeneous field, the function of that field being to make the position coordinate dependent on the original spin.

Let us call this apparently (but not really) random change in θ a “kick.” As θ varies, the number of kicks to one or the other special value also varies. Let $f(\theta)$ be the probability of obtaining a kick of size θ . Of course for any one experiment only kicks of size $-\theta + n\pi$ (with n an integer) enter, but it is reasonable to assume that as θ varies there is some well-defined distribution, manifested in our situation as a probability. Thus to get spin up one must have a kick of size $-\theta$ or, if one allows for larger kicks, $2n\pi - \theta$. Similarly, to get spin down one requires a kick of size $(2n+1)\pi - \theta$ (with n an integer). Thus the probability of spin up is $g(\theta) = \sum_{n=-\infty}^{n=\infty} f(-\theta + 2n\pi)$. Similarly for spin down one adds π to each summand in the argument of f . On the other hand, standard quantum mechanics, i.e., the Born rules, dictate that the ratio of down to up is $\tan^2(\theta/2)$. Therefore our requirement on f (hence on g) is

$$\tan^2\left(\frac{\theta}{2}\right) = \frac{g(\theta + \pi)}{g(\theta)}, \quad \text{with} \quad g(\theta) \equiv \sum_{n=-\infty}^{n=\infty} f(\theta + 2n\pi). \quad (5)$$

(In Eq. (5), in the definition of g , use has been made of f 's $\theta \rightarrow -\theta$ symmetry as well as the fact that n is a dummy variable.) There is an explicit solution to Eq. (5), namely $f_0(\theta) = 1/\theta^2$. Unfortunately this solution is not normalizable, as a probability should be. (In fact there is no normalizable solution to Eq. (5); see App. 2.) However, for θ close to 0 it is possible to cut off the function and eliminate the singularity without experimental implications. A convenient cutoff makes use of the Cauchy distribution

¹¹ In the companion article [10] we discuss this issue more thoroughly.

$$C_a(\phi) = \frac{a/\pi}{a^2 + \phi^2}, \quad (6)$$

which for small enough a changes Eq. (5) very little (see App. 3). The deviations from standard probabilities are largest for $\theta \sim 0$ and are of order a^2 ; since a is unknown, one does not have an experimental test. The distribution C_a does not have a second moment (it is a Lévy distribution), and this may be necessary for any function f satisfying Eq. (5). In App. 2 I report partial proofs, but I emphasize that for the purposes of the experimental tests described below and in [10], the Cauchy distribution is not required.

Before going to what I consider a true experimental test I will discuss two possibilities, one of which does not constitute a test, while the other may, but I don't know how to estimate outcomes.

The first—the non-test—deals with fluctuations. Perhaps with a Cauchy distribution at the heart of the special states, fluctuations would be exceptionally large. In other words, you send in N atoms with initial wave function given by Eq. (4) and the average number of spin down outcomes would be $N \sin^2(\theta/2)$; nevertheless, the *fluctuations* about that average would be larger than the expected \sqrt{N} . For better or for worse, this is not true, and a mathematical demonstration is given in App. 4.

A second test deals with runs, i.e., are the successive (in time) spin values, as detected by the (Stern-Gerlach) screen independent of one another, or do they tend to have many up, followed by many down, etc., keeping the average correct. Why should this be? If the special state depends on fluctuations in the magnetic fields, it is plausible that fewer unusual field values would be needed if successive “kicks” were correlated. This may indeed be happening, but I have no way to estimate it.

5 Force-Free Rotation?

In the MWI there is no need for any force to be applied in order to go from (say) an eigenfunction of J_x to an eigenfunction of J_z . The same is true for the Copenhagen interpretation.

What is amusing is that individual observers do have a change in their perceptions of the value of the angular momentum, and this occurs because of a change in the overall wave function, *but* there is no need for angular momentum non-conservation. The context is the Stern-Gerlach experiment measuring the z component of angular momentum. Imagine an (already) up spin sent into this apparatus. There is no transfer of J_z , although there is a very small transfer of linear momentum, since the atom (carrying the spin) is deflected. Similarly a down spin induces no J_z transfer. In equations

$$\begin{aligned} |\psi\rangle_+ &= |\uparrow\rangle \otimes \Omega_0 \otimes |\text{observer}\rangle \rightarrow U|\psi\rangle_+ = |\uparrow\rangle \otimes \Omega_+ \otimes |\text{observer sees } +\rangle \\ |\psi\rangle_- &= |\downarrow\rangle \otimes \Omega_0 \otimes |\text{observer}\rangle \rightarrow U|\psi\rangle_- = |\downarrow\rangle \otimes \Omega_- \otimes |\text{observer sees } -\rangle. \end{aligned} \quad (7)$$

In these equations $U = \exp(-iHt/\hbar)$ is the time evolution operator and Ω represents anything not the spin or the observer, in particular the magnets and the atom's translational degrees of freedom. I emphasize: although there is a small transfer of linear

momentum, *there no transfer of z -component of angular momentum in either case.* But now consider an initial state whose spin is not oriented along the z -axis,

$$\psi_{\text{initial}} = (\alpha|\uparrow\rangle + \beta|\downarrow\rangle) \otimes \Omega_0 \otimes |\text{observer}\rangle. \quad (8)$$

By the superposition principle this yields one observer who had prepared the state at a non-trivial angle to the z -axis, but found at the end a spin pointing along that axis. The same is true of the other observer. Each has seen a change in the perception of the z -component of angular momentum. On the other hand, since the wave function in Eq. (8) is a superposition of the initial wave functions of Eq. (7) the dynamics can be separately considered for each, and there is no transfer of z -component of angular momentum at any stage. How can this be? The (version of the) observer who saw “up” will say, “Oh, I was on the “ α ” (in our notation) component of the wave function, while the other (version of the) observer would make a similar statement, with β replacing α .

One should not find this shocking. Despite the observer’s possible perplexity, there *is* conservation of angular momentum. The total Hamiltonian commutes with the (total) J_z operator; it’s just that each observer, decohered from the other, sees a peculiarity.

The explanation would be slightly different for (my understanding of) the Copenhagen interpretation. Until you actually measure J_z it has no value, since J_z does not commute with the projector for the spin state in Eq. (8).

However, with only one world—the contention of the special state theory—there can be no change in the wave function without a proximate cause. If a quantity is changed the single observer can, if it is physically possible, determine what caused the change. This proximate cause lies in the special state itself. If $\langle |J_z| \rangle$ (of the spin) changes its value, something else has to pick it up.¹² This “something” can only be due to the peculiarities of the special state, what has been called a kick earlier. (Recall, the kick is not a deviation from the laws of nature, but like the cat at time-9 in the progression of Fig. 5, it is the result of exact obedience to the rules, but happening because of unusual initial conditions.)¹³

In a companion article [10] we give concrete suggestions for detecting the ostensibly random cause in some particular experiments. The general idea is to find where the unusual state (giving rise to the “kick”) is least unlikely, and attempting to detect it.

6 Discussion

After giving background on the special state theory I arrived at a potential experimental test. The many worlds and Copenhagen interpretations predict observations of source-free changes of the observer’s perception of the angular momentum. The special state theory, on the other hand, does not: something must push the spin to its new orientation. The proposed experimental test makes use of this distinction.

¹² The fact that $[H, J_z] = 0$ guarantees that the expectation of the operator does not change. In MWI this allows the two versions of the observer to balance $\langle |J_z| \rangle$, but this is not possible in the special state theory—there is only one observer at the end of the experiment.

¹³ I mention here *initial* conditions, but the conditions could be specified at any time, since the dynamics, both for the cat and for quantum mechanics, are completely deterministic.

In the appendices special topics related to the main text are taken up. In particular there is a proof that no probability distribution can exactly satisfy Eq. (5) and a demonstration of the preferred role of the Cauchy distribution—preferred even over other long-tailed distributions.

Acknowledgments I am grateful to Predrag Cvitanovic, Marcos da Luz, Bernard Gaveau, Amos Ori and Shina Tan for helpful discussions. Open access funding has been provided by the Max Planck Society.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix 1: One Way to Find Special States

For the purposes of the present article, I define “special” to mean the system is initially in some subspace of Hilbert space (\mathcal{H}) at time-0 and in a second subspace (possibly the same one) at a particular later time, T . In the spin-boson example of Sect. 2 the initial subspace was $\{\text{spin up}\} \otimes \{\text{all oscillator states}\}$. Let the projector for this subspace be called P . Let the final subspace have projector Q . For our spin-boson example, $Q = P$ or $Q = \text{projection operator for } \{\text{spin down}\} \otimes \{\text{all oscillator states}\}$, which is why it was appropriate to call the quantity in Fig. 2 the survival probability, survival in the subspace $P\mathcal{H}$. If the initial state $\psi \in P\mathcal{H}$, letting $U(t) \equiv \exp(-iHt/\hbar)$, then the amplitude in $Q\mathcal{H}$ at time- t is $A\psi$, with $A \equiv QU(t)P$. The probability of going from one subspace to the other is thus $\langle \psi | A^\dagger A \psi \rangle$, so that to say that a vector leaves one *entirely* in $P\mathcal{H}$ at time- t is to say that $B \equiv A^\dagger A$ has an eigenvalue 1. Finding the spin to be entirely down would mean that the overall state vector is in the subspace $Q\mathcal{H}$ with $Q = \text{projection operator for } \{\text{spin down}\} \otimes \{\text{all oscillator states}\}$, which is to say, the operator B has an eigenvalue 0. The search for special states thus takes the form of studying the spectrum of the operator B . This is what is done with respect to the model given in Eq. (1).

Appendix 2: The Functional Equation, Eq. (5)

Eq. (5) is

$$\tan^2\left(\frac{\theta}{2}\right) = \frac{g(\theta + \pi)}{g(\theta)}, \quad \text{with} \quad g(\theta) \equiv \sum_{n=-\infty}^{n=\infty} f(\theta + 2n\pi). \quad (9)$$

where $f(\theta)$ is a probability distribution, symmetric about 0 and monotonically decreasing as $|\theta|$ increases. As remarked, $f_0(\theta) = 1/\theta^2$ solves Eq. (9), but is not integrable and therefore cannot be a probability distribution. I first show that there is no continuous integrable solution.

Suppose the contrary. Define $h(\theta) \equiv g(\theta) \sin^2(\theta/2)$. By Eq. (9) h has period π . Now $f(0)$ must be (strictly) positive and therefore g is as well. Moreover, g is integrable on $[-\pi, \pi]$ since f is integrable on $[-\infty, \infty]$. This implies that $h(0) = 0$,

since otherwise g would not be integrable. By the periodicity of h , I must also have that $h(\pi) = 0$.¹⁴ But $\sin^2(\pi/2) = 1$, implying that $g(\pi) = 0$, which in turn implies that $f(\pi) = 0$ since f is non-negative. By monotonicity this implies that the entire support of f is on $[-\pi, \pi]$ and $f = g$. Let $\theta = \pi + \phi > \pi$. Then by the periodicity of h , $h(\theta) = h(\phi) = f(\pi + \phi) \sin^2(\pi + \phi) = 0 \times \text{a positive constant} = 0$, so h vanishes everywhere, which is impossible.

So the best that can be done is to solve the functional equation approximately. One can write

$$\begin{aligned}\Pr(\text{UP}) &= \cos^2(\theta/2) - s(\theta) \\ \Pr(\text{DN}) &= \sin^2(\theta/2) + s(\theta),\end{aligned}\quad (10)$$

with the assumption that $|s(\theta)| \ll 1$. This form is dictated by the demand that the probabilities add to 1. It is physically clear that $s(\theta)$ is real and symmetric about $\theta = 0$. Moreover, non-negativity of the probabilities demands that

$$\Pr(\text{UP}) \geq 0 \Rightarrow \cos^2(\theta/2) \geq s(\theta), \quad \Pr(\text{DN}) \geq 0 \Rightarrow \sin^2(\theta/2) \geq -s(\theta), \quad (11)$$

With this, the modified form of Eq. (9) is

$$\frac{\Pr(\text{DN})}{\Pr(\text{UP})} = \frac{\sin^2(\theta/2) + s(\theta)}{\cos^2(\theta/2) - s(\theta)} = \frac{g(\theta + \pi)}{g(\theta)}. \quad (12)$$

(Note that for small s the ratio becomes $\tan^2(\theta/2) + s(\theta)/\cos^4(\theta/2)$.) I again define

$$h(\theta) = g(\theta) \sin^2(\theta/2) \quad (13)$$

and Eq. (12) becomes

$$h(\theta + \pi) - h(\theta) = s(\theta) (g(\theta + \pi) + g(\theta)). \quad (14)$$

Dividing, I obtain an expression for the correction term, $s(\theta)$. It is then clear that it automatically satisfies the inequalities of Eq. (11). In particular it follows that

$$\begin{aligned}\cos^2(\theta/2) - s(\theta) &= \frac{g(\theta)}{g(\theta) + g(\theta + \pi)} \quad \text{and} \\ s(\theta) + \sin^2(\theta/2) &= \frac{g(\theta + \pi)}{g(\theta) + g(\theta + \pi)},\end{aligned}\quad (15)$$

both of which are non-negative. I next Fourier transform. Recall the definition and inversion:

$$\tilde{F}(n) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dx F(x) e^{-inx}, \quad F(x) = \sum_n \tilde{F}(n) e^{inx}. \quad (16)$$

¹⁴ It is also true that the symmetry of h about 0 (following from that of g) implies that for small θ , $h(\theta) = O(\theta^2)$, but we will not use this.

and the Poisson summation formula

$$\sum_{k=-\infty}^{\infty} F(x + 2\pi k) = \sum_{n=-\infty}^{\infty} \tilde{F}(n) e^{inx}, \quad (17)$$

It immediately follows that

$$g(\theta) = \sum_{n=-\infty}^{\infty} \tilde{f}(n) e^{in\theta}, \quad (18)$$

and

$$h(\theta) = \sin^2 \frac{\theta}{2} \sum_{n=-\infty}^{\infty} \tilde{f}_n e^{in\theta}. \quad (19)$$

Then

$$h(\theta + \pi) - h(\theta) = \sum_{n=-\infty}^{\infty} \tilde{f}_n e^{in\theta} \left(e^{in\pi} \cos^2 \frac{\theta}{2} - \sin^2 \frac{\theta}{2} \right) \quad (20)$$

$$= -\frac{1}{2} \sum_{n \text{ odd}, -\infty}^{\infty} e^{in\theta} (2\tilde{f}_n - \tilde{f}_{n+1} - \tilde{f}_{n-1}) \quad (21)$$

$$= s(\theta) (g(\theta + \pi) + g(\theta)) \quad \text{by Eq. (14)} \quad (22)$$

$$= 2s(\theta) \sum_{\text{even } n=-\infty}^{\infty} \tilde{f}(n) e^{in\theta}. \quad (23)$$

Comparing Eq. (21) and Eq. (23) we see that s can only have odd Fourier components. The reality of $s(\theta)$ implies that $\tilde{s}_n = \tilde{s}_{-n}^*$ while the symmetry implies $\tilde{s}_n = \tilde{s}_{-n}$, implying that \tilde{s}_n are real and symmetric with respect to n . As a result

$$s(\theta) = 2 \sum_{\text{odd } n=1}^{\infty} \tilde{s}(n) \cos(n\theta). \quad (24)$$

The first term in this expression is $\cos \theta$, in agreement with Eq. (34), the result of the Cauchy distribution.

Because $f(\theta)$ is real and symmetric about 0, its Fourier coefficients are necessarily real, showing that only cosine terms appear. Using Eqs. (14), (21) and (23), one obtains a formula for $s(\theta)$:

$$s(\theta) = \frac{\sum_{k \geq 0} \cos((2k+1)\theta) (\tilde{f}(2k+2) + \tilde{f}(2k) - 2\tilde{f}(2k+1))}{2(\tilde{f}(0) + 2 \sum_{k \geq 1} \cos(2k\theta) \tilde{f}(2k))}. \quad (25)$$

Next I review the arguments in [1] but with a clearer statement of assumptions and conclusions. The hypothesis that is refuted there is one that might reasonably come

to mind. Given that large kicks would be difficult to explain and perhaps would have been noticed, one might assume that the total kick that a given spin would be subject to would be the sum of many small kicks.¹⁵ If those “small kicks” have a finite second moment (and why shouldn’t they, if they’re small) then by the central limit theorem, the total kick, whose probability distribution is $f(\theta)$ would be a Gaussian, $f(\theta) = \exp(-\theta^2/2\mu^2)\sqrt{2\pi}\mu$ for some μ . From Eq. (16), $\tilde{f}(n) = (1/2\pi)\exp(-n^2\mu^2/2)$. Using Eq. (25) one obtains

$$s(\theta) = \frac{\sum_{k \geq 0} \cos((2k+1)\theta) e^{-(2k+1)^2\mu^2/2} \left\{ e^{-\mu^2/2} \cosh((2k+1)\mu^2) - 1 \right\}}{(1 + 2 \sum_{k \geq 1} \cos(2k\theta) e^{-2k^2\mu^2})}. \quad (26)$$

For $\mu = 0$, $s(\theta) = 0$, but the associated f is not an acceptable distribution and the limit $\mu \rightarrow 0$ is singular. Consider small μ , going back to the distribution in θ . Then $g(\theta) \approx f(\theta)$ since for small enough μ , the entire sum in g is dominated by its first term. In this case $\Pr(\text{UP}) = f(\theta)/[f(\theta) + f(\theta + \pi)]$. For θ , say 30° , this gives a number very close to one, not $\cos^2(15^\circ)$. On the other hand, for large μ , the only surviving term in Eq. (26) is $s = \cos(\theta)/2$, which is certainly not small. At intermediate values of μ numerical methods for Eq. (26) are reliable and s can be seen to differ significantly from zero.

The selection of the Cauchy distribution—even in preference to other long-tailed distributions not possessing a second moment—can be seen in the following way. One expects f to be very steeply peaked around 0, since most microscopic states do nothing to the spins. So it is reasonable to take

$$f(\theta) = \frac{1}{ac_\ell} \left(\frac{a^2}{a^2 + \theta^2} \right)^\ell, \quad \text{with} \quad c_\ell = \int_{-\infty}^{\infty} \frac{du}{(1+u^2)^\ell} = \frac{\Gamma(1/2)\Gamma(\ell-1/2)}{\Gamma(\ell)} \quad (27)$$

taking care of the normalization.¹⁶ With this ansatz for f one can numerically compute the terms in Eq. (14) and divide in order to arrive at $s(\theta)$. I have done this for a number of ℓ values both near 1 and near 1.5, where the latter is the minimum value for the existence of a second moment. I first show various measures of the size of $s(\theta)$. In Fig. 7 are shown 3 quantities for each ℓ value. They are the maximum of $|s(\theta)|$, its mean and its standard deviation. In this logarithmic plot it is clear that $\ell = 1$, the Cauchy distribution, is vastly superior to all the others. This assertion includes the

¹⁵ “Small kicks” would need to be separated by the physical correlation time in order to be considered independent.

¹⁶ There is little loss of generality in using this form for $f(\theta)$. For all cases, we expect a to be small, so that the tail of the distribution function has the correct properties. For $\ell < 3/2$ the asymptotic behavior in θ is $f \sim 1/\theta^{2\ell}$ so that the only difference from the Lévy distribution is the small θ behavior. (For $\ell = 1$ this is precisely the Cauchy distribution.) For $\ell \geq 3/2$ one can use the central limit theorem and one goes back to the earlier discussion in the text. The only case not considered is large a and $\ell \geq 3/2$. Numerical investigation of this case found a minimum in size for $s(\theta)$ at $a = 1$, for a variety of ℓ values. This minimum was about e^{-4} in sharp contrast to the much smaller values for the Cauchy distribution using small a .

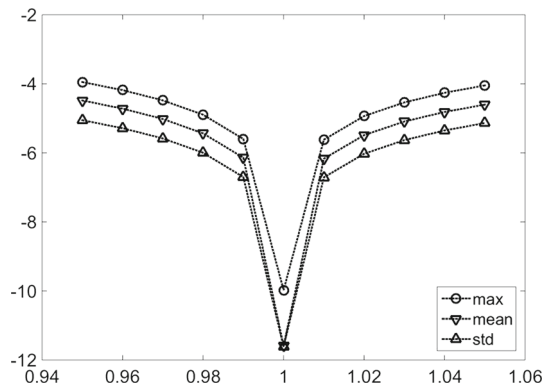


Fig. 7 Size of the function $s(\theta)$ for various values of ℓ . The ordinate is the natural logarithm of various measures of the size of $s(\theta)$. The value $\ell = 1$, the Cauchy distribution, clearly stands out as optimal. Three measures are used, the maximum of $|s(\theta)|$ (as a function of θ) and the mean and standard deviation of that absolute value. I do not bother to put a detailed calculation of the neighborhood of $\ell = 1.5$ since the values of all measures of s are even larger than those shown in this figure. (What is special about 1.5 is that that is the threshold for the existence of a second moment. Thus for all values given in the above figure there is no second moment. For the curves in this figure $a = 0.01$).

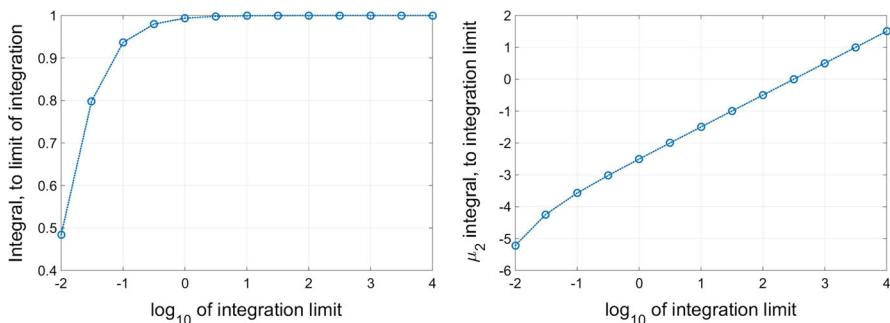


Fig. 8 Slow convergence of the analytically known infinite second moment. For $a = 0.01$, adding terms for $|x| < 1$ already provides most of the probability. But it takes 10^4 terms for the second moment to even reach 10.

neighborhood of $\ell = 3/2$ where the second moment begins to exist. Near $\ell = 1.5$ I do not bother to show a graph since even the smallest measure of s 's size (the standard deviation) has a (natural) logarithm of -3 , in contrast to values near $\ell = 1$, where values from -6 to -12 prevail.

It must be said that the numerical calculations performed here require considerable care. The problem is the slow convergence of the sums involved. As an illustration I first look at the second moment for the Cauchy distribution. This is analytically known to be infinite. However, in Fig. 8 I show how the computed value of this quantity grows as more and more terms are included. For $a = 1/100$ the integral of the probability distribution is nearly unity when you include all terms for $|x| \leq 1$. However, at that point the value of the second moment is only about 0.003. Note that

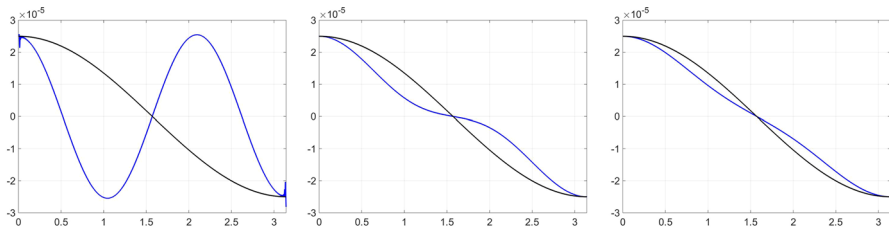
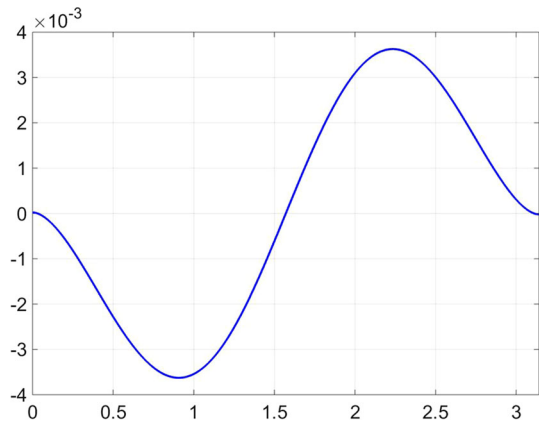


Fig. 9 Comparison of the numerically calculated $s(\theta)$ with the analytic value given in Eq. (34). The numerical value depends on the function g which is approximated by $\Delta\theta \sum_n \sum_{k=-N}^N f(\theta_n + 2k\pi)$, for $\theta_n = -n\Delta\theta$ such that $n \in [-R/\Delta\theta, +R/\Delta\theta]$. For the three plots $\Delta\theta = 0.01, 0.0025, 0.0001$, $N = 500, 2500, 5000$ and $R = 500, 1500, 1500$. Note that despite the poor shape fit for the least accurate calculation, the correct magnitude of $s(\theta)$ is obtained.

Fig. 10 A slight shift in ℓ away from its Cauchy distribution value enlarges $s(\theta)$ tremendously. Here I show a graph of s for $\ell = 1.01$. To the eye, the associated g is nearly indistinguishable from its $\ell = 1$ value; nevertheless, $s(\theta)$ is sensitive to this small change and increases significantly over its $\ell = 1$ value.



the graph of the second moment is logarithmic and it takes $|x| \sim 10^4$ terms to even bring the second moment over 10. Nevertheless, it does grow to infinity. Corresponding issues affect the calculation of $g(\theta)$. Many consistency checks were therefore made but even so one can see discrepancies. An example for the need of extreme accuracy is shown in Fig. 9. There I compute $s(\theta)$ for $\ell = 1$ at three levels of numerical precision. In the worst case, the range of θ was taken to be $[-500, 500]$, the mesh (the numerical “ $\Delta\theta$ ”) was 0.01 and the number of terms added to obtain $g(\theta)$ was 1000. The graph of g under these circumstances is not bad—nevertheless when used to compute s and compared to the known analytic result (Eq. (34)) it is evidently poor. By contrast in the most accurate calculation that I performed the fit begins to look good. In this case 10,000 terms were added in the computation of g , the mesh was 10^{-4} and the range was not 500, but 1500. Note though that despite the variation in shape, for all three calculations, the magnitude of $s(\theta)$ was about the same, bounded by 10^{-5} . By contrast (Fig. 10), for the nearby ℓ value, 1.01, the computed $s(\theta)$ is dismal: not only is the shape wrong, but the magnitude scale is already significantly larger.

Appendix 3: Properties of the Cauchy Distribution.

The Cauchy distribution is

$$C_a(\phi) = \frac{a/\pi}{a^2 + \phi^2}. \quad (28)$$

It is a probability distribution having integral 1. It *almost* satisfies Eq. (5), and as such is a form of cutoff distribution solving that equation. It is a Lévy distribution and has many properties different from the standard normal (or Gaussian) distribution. In this appendix I gather properties of this function.

The basic identity comes from analytic function theory see § 9.3, Eq. 9.3.14. [11]:

$$\frac{1}{\tan z} = \sum_{n=-\infty}^{\infty} \frac{1}{z - n\pi}, \quad (29)$$

where n runs over the integers. The poles of one over the tangent function occur at multiples of π and the residues are unity. Let $z = \psi + ib$. Using elementary relations I write the real and imaginary parts of Eq. (29),

$$\frac{\tan \psi}{\tan^2 \psi \cosh^2 b + \sinh^2 b} = \sum_n \frac{\psi - n\pi}{(\psi - n\pi)^2 + b^2}, \quad (30)$$

$$\frac{\tanh b}{\tanh^2 b \cos^2 \psi + \sin^2 \psi} = \sum_n \frac{b}{(\psi - n\pi)^2 + b^2}. \quad (31)$$

With $f(\theta) = C_a(\theta)$, $\psi = \theta/2$ and $b = a/2$, this implies that

$$g_C(\theta) = \sum_{k=-\infty}^{\infty} C_a(\theta + 2k\pi) = \frac{1}{2\pi} \frac{\tanh(a/2)}{\tanh^2(a/2) \cos^2(\theta/2) + \sin^2(\theta/2)}, \quad (32)$$

where g_C is the “ g ” associated with the Cauchy distribution. It follows that

$$\begin{aligned} \frac{g_C(\theta + \pi)}{g_C(\theta)} &= \frac{\tanh^2(a/2) \cos^2(\theta/2) + \sin^2(\theta/2)}{\tanh^2(a/2) \sin^2(\theta/2) + \cos^2(\theta/2)} \\ &= \frac{\tanh^2(a/2) + \tan^2(\theta/2)}{\tanh^2(a/2) \tan^2(\theta/2) + 1}. \end{aligned} \quad (33)$$

This obviously goes to the correct limit as $a \rightarrow 0$ and just as obviously does not for finite a . It is also possible to calculate the deviations from standard probabilities. The probability of finding spin up is

$$\begin{aligned} \text{Pr(UP)} &= \frac{\text{Prob. for special kicks UP}}{\text{Prob. for all special kicks}} = \frac{g_C(\theta)}{g_C(\theta) + g_C(\theta + \pi)} \\ &\approx \cos^2(\theta/2) - \frac{a^2}{4} \cos \theta. \end{aligned} \quad (34)$$

This differs from the standard value, and in principle could be used to place a bound on a . But this would not be an experimental test unless experiment showed a non-zero value of a .

By Eq. (16) the Fourier coefficients of C_a are

$$\tilde{C}_a(n) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dx \frac{a/\pi}{a^2 + \phi^2} e^{-in\phi} = \frac{1}{2\pi} \exp(-a|n|) \quad (35)$$

If one sums over the even components only (as in Eq. (23)), for the Cauchy distribution these are simple:

$$\sum_{\text{even } n=-\infty}^{\infty} \tilde{C}_a(n) \exp(-in\theta) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \exp(-2a|k| - ik2\theta) = C_{2a}(2\theta). \quad (36)$$

Integer Powers of the Cauchy Distribution

For reference and for the ability to make numerical checks, I study powers of the Cauchy distribution. First the square:

$$C_a^{(2)}(\phi) = \frac{2a^3}{\pi} \left(\frac{1}{a^2 + \phi^2} \right)^2. \quad (37)$$

It is a probability distribution having integral 1. Its second moment is

$$\mu_2 = \int_{-\infty}^{\infty} dx x^2 C_a^{(2)}(x) = a^2. \quad (38)$$

Its Fourier coefficients are

$$\tilde{C}_a^{(2)}(k) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dx \exp(-ikx) C_a^{(2)}(x) = \frac{1}{2\pi} (a|k| + 1) e^{-a|k|}. \quad (39)$$

The third power is

$$C_a^{(3)}(\phi) = \frac{8a^5}{3\pi} \left(\frac{1}{a^2 + \phi^2} \right)^3. \quad (40)$$

It is a probability distribution having integral 1. Its second moment is

$$\mu_2 = \int_{-\infty}^{\infty} dx x^2 C_a^{(3)}(x) = \frac{a^2}{3}. \quad (41)$$

Its Fourier coefficients are

$$\tilde{C}_a^{(3)}(k) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dx \exp(-ikx) C_a^{(3)}(x) = \frac{1}{2\pi} \frac{(a^2|k|^2 + 3a|k| + 3)}{3} e^{-a|k|}. \quad (42)$$

Results of this subsection were used as analytic tests of numerical results. As mentioned in Sect. 2, convergence of many sums is extremely slow and numerical results must be carefully checked. Although μ_2 converges nicely for Eq. (37), μ_4 does not, and for the Fourier coefficients the analytic results provide bona fide checks of numerical accuracy.

Appendix 4: A Statistical Measure that is *not* an Experimental Test

As discussed above, in (our version of¹⁷) the Stern-Gerlach experiment you send in a beam of atoms prepared so that for all of them the spin degree of freedom has the wave function $u_\theta = e^{i\theta\sigma_x/2} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. At the end of the experiment—after the atoms have passed through an appropriate non-uniform magnetic field—the spins are found to be either spin-up (implying a wave function of the form $(\exp(i\omega), 0)^\dagger$ (ω real)) or spin-down. The analysis considers “kicks” that align the spins “up” or “down,” and demands that the underlying distribution of these kicks yield $\tan^2(\theta/2)$ as the ratio of down to up spins. A distribution accomplishing this does not have a second moment, and would be the Cauchy distribution with small parameter “ a ,” to wit

$$C_a(\phi) = \frac{a/\pi}{a^2 + \phi^2}. \quad (43)$$

The deviation from the standard (\tan^2) result is $O(a^2)$. One might propose the following experimental test. Since the Cauchy distribution has such a long tail, i.e., drops off so slowly in ϕ that it possesses no second moment, does not self-average, and possesses other differences from the well behaved normal distribution, perhaps the counting statistics in the Stern-Gerlach experiment will reflect this.

Let me be more specific. Consider the permissible outcomes of the Stern-Gerlach experiment: The value of the angle of the kick, $\omega = -\theta + \phi$, must either fall in the

¹⁷ To prepare the atoms in the state shown in Eq. (4) one could use half the output of another Stern-Gerlach setup oriented at an appropriate angle. Something like this was done by Stern himself in the early 1930's. The problem he and his collaborators came up against was that if there was a non-zero magnetic field along the entire path between the two magnets the spins would follow the field. Stern et al. summarized their work in [15], but it is interesting to read their tribulations in the course of these efforts [16, 17]. The analysis of their work was done analytically by Majorana [18]. I have translated both the Frisch-Segre article (from the German) and the Majorana article (from the Italian) (however, imperfectly) and would be happy to send this to anyone who asks (and I'd be grateful for corrections). The Majorana calculation was a tour de force and was sufficient to allow Frisch and Segre to conclude that theory and experiment were consistent. It is amusing that in the 21st century I could use a computer to take into account effects that Majorana neglected and found even better agreement of theory and experiment. My translations and computations are unpublished.

interval $[-b/2, b/2]$ or $[\pi - b/2, \pi + b/2]$ for small b . The probability of each outcome can be evaluated and in fact gives the ratio $\tan^2(\theta/2)$. (So the Cauchy distribution works.) If you send in N atoms, the *average* number of down spins will be $\sin^2(\theta/2)N$ (and similarly for up), but the question is, are the deviations large?

The answer is “no,” as I now show. Let $\{\Phi_k\}$, $k = 1, \dots, N$ be independent identically distributed random variables having the Cauchy distribution. For given θ , the number that land in the interval $[-b/2, b/2]$, call it n_b , is

$$n_b = \sum_{k=1}^N \int_{-b/2}^{b/2} d\phi_k \sum_{\ell=-\infty}^{\infty} \Theta\left(\frac{b}{2} - |\phi_k + 2\pi\ell - \theta|\right), \quad (44)$$

with Θ the Heaviside function, ϕ_k the value taken by Φ_k , and ℓ running over the integers. For any one Φ_k the associated probability can be evaluated if one takes b sufficiently small, namely (calling p that probability)

$$p = b \frac{a}{\pi} \sum_{\ell=-\infty}^{\infty} \frac{1}{a^2 + (\theta + 2\pi\ell)^2} = \text{prob. that the random variable falls in } \left[-\frac{b}{2}, \frac{b}{2}\right]. \quad (45)$$

This sum can be evaluated to yield

$$p = \frac{(b/2\pi) \tanh(a/2)}{\sin^2(\theta/2) + \cos^2(\theta/2) \tanh^2(a/2)}. \quad (46)$$

Now I want the probability that exactly n such random variables land in the given intervals. Since the order in which they appear does not matter this is given by

$$\text{Prob}\left(n = \sum_{k=1}^N \text{True}(\Phi_k \in [-b/2, b/2] \bmod 2\pi)\right) = \binom{N}{n} p^n (1-p)^{N-n}, \quad (47)$$

with “True” the function that takes the value 1 when its argument is true, 0, otherwise. As a function of n , the probability distribution in Eq. (47) is known to be well-behaved, to have a second moment and similar good properties. The underlying Cauchy distribution does not impair this distribution. The only trouble could arise when θ is small compared to a , but the dependence on the value of a precludes an experimental test of this behavior. (b too must be sufficiently small, but that is not a problem.)

Note that in this demonstration I made essential use of the independence of the “kicks.” The existence of correlations, as contemplated at the end of Sect. 4 would invalidate this argument.

References

1. Schulman, L.S.: Time's Arrows and Quantum Measurement. Cambridge University Press, New York (1997)
2. Schulman, L.S.: Special states in the Spin-Boson model. *J. Stat. Phys.* **77**, 931–944 (1994)
3. Gaveau, B., Schulman, L.S.: Is ergodicity a reasonable hypothesis? *Eur. Phys. J. Spec. Top.* **224**, 891–904 (2015)
4. Arnol'd, V.I., Avez, A.: Ergodic Problems of Classical Mechanics. Benjamin, New York (1968)
5. Wald, R.M.: The arrow of time and the initial conditions of the universe. [arXiv:gr-qc/0507094v1](https://arxiv.org/abs/gr-qc/0507094v1) (2005)
6. Khoury, J., Steinhardt, P.J., Turok, N.: Designing cyclic universe models. *Phys. Rev. Lett.* **92**, 031302 (2004)
7. Steinhardt, P.J., Turok, N.: Cosmic evolution in a cyclic universe. *Phys. Rev. D.* **65**, 126003 (2002)
8. Gold, T.: The arrow of time. *Am. J. Phys.* **30**, 403–410 (1962)
9. Schulman, L.S.: Evolution of wave-packet spread under sequential scattering of particles of unequal mass. *Phys. Rev. Lett.* **92**, 210404 (2004)
10. Schulman, L.S., da Luz, M.G.E.: Looking for the source of change. *Found. Phys.* (2016). doi:[10.1007/s10701-016-0031-x](https://doi.org/10.1007/s10701-016-0031-x)
11. Hille, E.: Analytic Function Theory. Ginn and Company, New York (1959)
12. Schulman, L.S.: Models for intermediate time dynamics with two-time boundary conditions. *Phys. A.* **177**, 373–380 (1991)
13. Broderick, A.E., Narayan, R., Kormendy, J., Perlman, E.S., Rieke, M.J., Doeleman, S.S.: The event horizon of M87. *Astrophys. J.* **805**, 179 (2015). doi:[10.1088/0004-637X/805/2/179](https://doi.org/10.1088/0004-637X/805/2/179)
14. Almheiri, A., Marolf, D., Polchinski, J., Sully, J.: Black holes: complementarity or firewalls? *J. High. Energ. Phys.* **02**, 062 (2013)
15. Frisch, R., Phipps, T.E., Segre, E., Stern, O.: Process of space quantisation. *Nature.* **130**, 892–893 (1932)
16. Phipps, T.E., Stern, O.: Über die Einstellung der Richtungsquantelung. *Z. Phys.* **73**, 185–191 (1932)
17. Frisch, R., Segre, E.: On the process of space quantization. II. *Z. Phys.* **80**, 610–616 (1933)
18. Majorana, E.: Atoms in an oriented, variable magnetic field. *Nuov. Cim.* **2**, 43–50 (1932)